ON A PROBLEM OF ERDŐS INVOLVING THE LARGEST PRIME FACTOR OF n

Aleksandar Ivić

Abstract. Let $P(n)$ denote the largest prime factor of an integer n (≥ 2), and let $N(x)$ denote the number of natural numbers n such that $2 \le n \le x$, and n does not divide $P(n)!$. We prove that

$$N(x) = x\left(2 + O\left(\sqrt{\log_2 x/\log x}\right)\right)\int_2^x \rho(\log x/\log t)\frac{\log t}{t^2}\,\mathrm{d}t,$$

where $\rho(u)$ is the Dickman-de Bruijn function. In terms of elementary functions we have

$$N(x) = x\exp\left\{-\sqrt{2\log x\log_2 x}\,(1 + O(\log_3 x/\log_2 x))\right\},$$

thereby sharpening and correcting recent results of K. Ford and J.-M. De Koninck and N. Doyon.

1. Introduction and statement of results

Let $P(n)$ denote the largest prime factor of an integer n (≥ 2). In 1991 P. Erdős [5] proposed the following problem: prove that $N(x) = o(x)$ $(x \to \infty)$, where $N(x)$ denotes the number of natural numbers n such that $2 \le n \le x$, and n does not divide $P(n)!$. This problem is connected to the so-called Smarandache function $S(n)$, the smallest integer k such that $n|k!$.

Erdős's assertion was shown to be true by I. Kastanas [13], and S. Akbik [1] proved later that $N(x) \ll x\exp(-\frac14\sqrt{\log x}\,)$ holds. K. Ford [8] proposed an asymptotic formula for $N(x)$. His Theorem 1 states that

$$N(x) \sim \frac{\sqrt{\pi}(1+\log 2)}{2^{3/4}}(\log x\log_2 x)^{3/4}x^{1-1/u_0}\rho(u_0) \qquad (x\to\infty). \tag{1.1}$$

In this formula one should correct the constant, as will be show in Section 4. Here, as usual, $\log_1 x \equiv \log x$ and $\log_k x = \log(\log_{k-1} x)$ for $k \ge 2$. The function $\rho(u)$ is

1991 *Mathematics Subject Classification.* 11N25, 11N37.

Key words and phrases. The largest prime factor of n, Dickman-de Bruijn function.

the continuous solution to the difference delay equation $u\rho'(u) = -\rho(u-1)$ with the initial condition $\rho(u) = 1$ for $0 \le u \le 1$ and is commonly called the Dickman (or Dickman-de Bruijn) function. The function $u_0 = u_0(x)$ is implicitly defined by the equation

$$\log x = u_0 \left(x^{1/u_0^2} - 1 \right),$$

so that

$$u_0(x) = \left(\frac{2\log x}{\log_2 x}\right)^{1/2} \left\{ 1 - \frac{\log_3 x}{2\log_2 x} + \frac{\log 2}{2\log_2 x} + O\left(\left(\frac{\log_3 x}{\log_2 x}\right)^2\right)\right\}. \tag{1.2}$$

The asymptotic formula (1.1) and the explicit expression (1.2) enabled Ford (op. cit.) in Corollary 2 to deduce that

$$N(x) = x\exp\left\{-(\sqrt{2}+o(1))\sqrt{\log x \log_2 x}\right\} \qquad (x \to \infty). \tag{1.3}$$

Unaware of Ford's work, J.-M. De Koninck and N. Doyon [3] recently tackled this problem again. In [3] they published the result that

$$N(x) = x\exp\left\{-(2+o(1))\sqrt{\log x \log_2 x}\right\} \qquad (x \to \infty). \tag{1.4}$$

Unfortunately, (1.4) is not true (it contradicts (1.3)). Namely the argument in [3] leading to their (3.8) does not hold, and the crucial parameter $u = \log x/\log y$ in the formula for $\Psi(x, y)$ has to be evaluated more carefully. A straightforward modification of their proof leads then again to (1.3).

The aim of this note is to sharpen (1.1) and (1.3), and to provide asymptotic formulas for sums of $S^r(n)$ and $1/S^r(n)$ when $r \in \mathbb{R}$ is fixed. The results are contained in the following

THEOREM 1.

$$N(x) = x\exp\left\{-\sqrt{2}L(x)\left(1 + g_0(x) + O\left(\left(\frac{\log_3 x}{\log_2 x}\right)^3\right)\right)\right\}, \tag{1.5}$$

where

$$L = L(x) = \sqrt{\log x \log\log x}, \tag{1.6}$$

and, for $r > -1$,

$$\begin{aligned} g_r(x) = {} & \frac{\log_3 x + \log(1+r) - 2 - \log 2}{2\log_2 x}\left(1 + \frac{2}{\log_3 x}\right) \\ & - \frac{(\log_3 x + \log(1+r) - \log 2)^2}{8\log_2^2 x}, \end{aligned} \tag{1.7}$$

THEOREM 2.

$$N(x) = x\left(2 + O\left(\sqrt{\frac{\log_2 x}{\log x}}\right)\right)\int_2^x \rho\left(\frac{\log x}{\log t}\right)\frac{\log t}{t^2}\,\mathrm{d}t. \tag{1.8}$$

THEOREM 3. *Let $r > 0$ be a fixed number. Then we have*

$$\sum_{2\le n\le x}\frac{1}{S^r(n)} = x\exp\left\{-\sqrt{2r}L(x)\left(1 + g_{r-1}(x) + O\left(\left(\frac{\log_3 x}{\log_2 x}\right)^2\right)\right)\right\}, \tag{1.9}$$

where $L(x)$ and $g_r(x)$ are given by (1.6) *and* (1.7), *respectively. Moreover, for any fixed integer $J \ge 1$ there exist computable constants $a_{1,r} = \zeta(r+1)/(r+1)$, $a_{2,r}\ldots, a_{J,r}$ such that*

$$\sum_{2\le n\le x} S^r(n) = x^{r+1}\sum_{j=1}^{J}\frac{a_{j,r}}{\log^j x} + O\left(\frac{x^{r+1}}{\log^{J+1} x}\right). \tag{1.10}$$

The asymptotic formula (1.8) is sharper than (1.5)–(1.7), since it is a true asymptotic formula, while (1.5)–(1.7) is actually an asymptotic formula for $\log\frac{N(x)}{x}$. On the other hand, the right-hand side of (1.5) is given in terms of elementary functions, while (1.8) contains the (non-elementary) function $\rho(u)$. Thus is seemed in place to give both (1.5) and (1.8), especially since the proof of (1.5) requires only the more elementary analysis which follows the work of De Koninck–Doyon [3].

The asymptotic formula (1.9) seems to be new, while (1.10) sharpens the results of S. Finch [7], who obtained (1.10) in the case $J = 1$.

Acknowledgement. I thank Prof. G. Tenenbaum for pointing out formula (2.6) to me, and indicating how it improves the error term $O(\log_3 x/\log_2 x)$ in the first version of Th. 2 (coming from (2.5)), to $O(\sqrt{\log_2 x/\log x})$. I also thank the referee for valuable remarks.

2. The necessary results of $\Psi(x, y)$

The proofs of both Theorem 1 and Theorem 2 depend on results on the function

$$\Psi(x, y) = \sum_{n\le x, P(n)\le y} 1 \qquad (2 \le y \le x), \tag{2.1}$$

which represents the number of n not exceeding x, all of whose prime factors do not exceed y. All the results that follow can be found e.g., in [9], [10] and [14]. We have the elementary bound

$$\Psi(x,y) \ll x\exp\left(-\frac{\log x}{2\log y}\right) \qquad (2\le y\le x) \tag{2.2}$$

and the more refined asymptotic formula

$$\begin{gathered}\Psi(x,y) = x\rho(u)\left(1+O\left(\frac{\log(u+1)}{\log y}\right)\right)\\ u=\log x/\log y,\ \exp((\log_2 x)^{5/3+\varepsilon})\le y\le x.\end{gathered} \tag{2.3}$$

Note that the Dickman–de Bruijn function $\rho(u)$ admits an asymptotic expansion, as $u\to\infty$, which in a simplified form reads

$$\rho(u) = \exp\left\{-u\left(\log u+\log_2 u - 1 + O\left(\frac{\log_2 u}{\log u}\right)\right)\right\}.$$

Let, for given $u>1$, $\xi=\xi(u)$ be defined by $u\xi = \mathrm{e}^{\xi}-1$, so that

$$\xi(u) = \log u + \log_2 u + \frac{\log_2 u}{\log u} + O\left(\left(\frac{\log_2 u}{\log u}\right)^2\right). \tag{2.4}$$

If $u>2$, $|v|\le \frac12 u$, then we have the asymptotic formula

$$\rho(u-v) = \rho(u)\exp\left\{v\xi(u)+O((1+v^2)/u)\right\}. \tag{2.5}$$

This result (see e.g., Hildebrand–Tenenbaum [10]) for bounded $|v|$ (see (3.3.9) of de la Bretèche–Tenenbaum [2]) yields the asymptotic formula

$$\rho(u-v) = \rho(u)\mathrm{e}^{v\xi(u)}\cdot\left\{1+O\left(\frac1u\right)\right\} \qquad (|v|\le v_0). \tag{2.6}$$

3. Proof of Theorem 1

We turn now first to the proof of Theorem 1. The idea is to obtain upper and lower bounds of the form given by (1.5)–(1.7). For the former, note that if $P^2(n)|n$, then n is counted by $N(x)$, since $P^2(n)$ cannot divide $P(n)!$. Therefore, if $T_0(x)$ denotes the number of natural numbers $n\le x$ such that $P^2(n)|n$, then $N(x)\ge T_0(x)$. The desired lower bound follows then from the formula for $T_0(x)$, which is a special case of the formula for ($r\ge0$ is a given constant)

$$T_r(x) := \sum_{2\le n\le x, P^2(n)|n}\frac{1}{P^r(n)}, \tag{3.1}$$

obtained by C. Pomerance and the author [12], and sharpened by the author in [11]. Namely for $T_r(x)$ it was shown that

$$T_r(x) = x\exp\left\{-(2r+2)^{1/2}L(x)\left(1+g_r(x)+O\left(\left(\frac{\log_3 x}{\log_2 x}\right)^3\right)\right)\right\}, \tag{3.2}$$

where $L(x)$ and $g_r(x)$ are given by (1.6) and (1.7). We note that Erdős, Pomerance and the author in [6] investigated the related problem of the sum of reciprocals of $P(n)$. They proved that (this result is also sharpened in [11])

$$\sum_{2\le n\le x}\frac{1}{P(n)} = \sum_{p\le x}\frac{1}{p}\Psi\left(\frac{x}{p},p\right) = x\left(1+O\left(\sqrt{\frac{\log_2 x}{\log x}}\right)\right)\int_2^x \rho\left(\frac{\log x}{\log t}\right)\frac{1}{t^2}\,\mathrm{d}t. \tag{3.3}$$

The same argument leads without difficulty to

$$T_0(x) = \sum_{p\le x}\Psi\left(\frac{x}{p^2},p\right) = x\left(1+O\left(\sqrt{\frac{\log_2 x}{\log x}}\right)\right)\int_2^x \rho\left(\frac{\log x}{\log t}\right)\frac{\log t}{t^2}\,\mathrm{d}t. \tag{3.4}$$

It remains yet to deal with the upper bound for $N(x)$. As in [3], we shall use the inequality (p denotes primes)

$$N(x) \le \sum_{2\le r\le \log x/\log 2}\ \sum_{p\le x^{1/r}}\Psi\left(\frac{x}{p^r},pr\right). \tag{3.5}$$

To see that (3.5) holds note that, if n is counted by $N(x)$, then n must be divisible by p^r which does not divide $P(n)!$. The condition $r\ge 2$ is necessary, since squarefree numbers q (>1) divide $P(q)!$. Moreover, $pr\le P(n)$ cannot hold, since otherwise the numbers $p, 2p, \ldots, rp$ would all divide $P(n)!$, and so would p^r, which is contrary to our assumption. Thus $n = p^r m\,(r\ge 2)$, $P(m)\le P(n) < pr$, which easily gives (3.5).

We restrict now the ranges of r and p on the right-hand of (3.5) to the ones which will yield the largest contribution.

The contribution of $r > 3L$ is trivially

$$\ll \sum_{r>3L}\sum_p xp^{-r} \ll x2^{-3L} = x\exp(-3\log 2\cdot L),$$

which is negligible, since $3\log 2 > \sqrt{2}$. Similarly, the contribution of $p > \exp(2L)$ is negligible, and also for $p > \exp((\log x)^{1/3})$ and $r > (\log x)^{1/6}\log_2 x$ the contribution is, for some $C > 0$,

$$\ll \sum_{r>(\log x)^{1/6}\log_2 x}\ \sum_{p>\exp((\log x)^{1/3})} xp^{-r} \ll x\exp(-C\sqrt{\log x}\cdot\log_2 x),$$

which is negligible.

With (3.5) and (2.3) we see that in the range

$$L_1 := \exp((\log x)^{1/3}) < p \le \exp(\tfrac14 L), \qquad r \le (\log x)^{1/6}\log_2 x$$

we have

$$\begin{aligned} u &= \frac{\log(x/p^r)}{\log pr} = \frac{\log x - r\log p}{\log p + \log r} \\ &= \frac{\log x + O((\log x)^{2/3+\varepsilon})}{\log p + O(\log_2 x)} = \frac{\log x}{\log p}\left(1 + O\left((\log x)^{\varepsilon - 1/3}\right)\right). \end{aligned}$$

Therefore by (2.3) we obtain, as $x \to \infty$,

$$\Psi\left(\frac{x}{p^r}, pr\right) \le \frac{x}{p^r}\exp\left(-(1+o(1))\frac{\log x}{\log p}\log\left(\frac{\log x}{\log p}\right)\right),$$

which gives

$$\begin{aligned} &\sum_{L_1 < p \le \exp(\frac14 L)} \sum_{r\ge 2} \Psi\left(\frac{x}{p^r}, pr\right) \\ &\ll x \sum_{L_1 < p \le \exp(\frac14 L)} \frac{1}{p^2} \max_{L_1 < p \le \exp(\frac14 L)} \exp\left(-(1+o(1))\frac{\log x}{\log p}\log\left(\frac{\log x}{\log p}\right)\right) \\ &\ll x\exp(-(2+o(1))L), \end{aligned}$$

which is negligible. The contribution of $p \le L_1, r > (\log x)^{1/6}\log_2 x$ is easily seen to be also negligible. This means that the main contribution to the right-hand side of (3.5) comes from the range

$$\exp(\tfrac14 L) \le p \le \exp(2L), \qquad 2 \le r \le (\log x)^{1/6}\log_2 x, \tag{3.6}$$

or more precisely, for some $C > \sqrt{2}$,

$$\begin{aligned} N(x) &\ll x\exp(-CL) \\ &\quad + \sum_{2\le r\le (\log x)^{1/6}\log_2 x} \; \sum_{\exp(\frac14 L)\le p\le \exp(2L)} \Psi\left(\frac{x}{p^r}, pr\right). \end{aligned} \tag{3.7}$$

By trivial estimation it transpires that the contribution of $r \ge 8$ in (3.7) will be negligible, so that we may write

$$N(x) \ll x\exp(-CL) + L\max_{c>0}\sum_{2\le r\le 7} \mathcal{T}_{c,r}(x), \tag{3.8}$$

where, for a given constant $c > 0$ and $r \in \mathbb{N}$,

$$\mathcal{T}_{c,r}(x) := \sum_{\exp(cL-1)<p\le\exp(cL)} \Psi\left(\frac{x}{p^r}, pr\right),$$

since (again by (2.3)) the contribution of $c \ge c_0$, c_0 a large positive constant, is negligible. If p is in the range indicated by $\mathcal{T}_{c,r}(x)$, then

$$u = \frac{\log xp^{-r}}{\log pr} = \frac{\log x}{cL}\left(1 + O\left(\sqrt{\frac{\log_2 x}{\log x}}\right)\right),$$

$$\log u = \log_2 x - \log c - \log L + O\left(\sqrt{\frac{\log_2 x}{\log x}}\right).$$

Therefore, for $2 \le r \le 7$,

$$\begin{aligned}\mathcal{T}_{c,r}(x) \ll \sum_{\exp(cL-1)<p\le\exp(cL)} \frac{x}{p^r}\exp\Big\{-\frac{\log x}{cL}\Big(1 + O\Big(\sqrt{\frac{\log_2 x}{\log x}}\Big)\Big)\times \\ \times\left(\tfrac{1}{2}\log_2 x + \tfrac{1}{2}\log_3 x + O(1)\right)\Big\} \\ \ll x\exp\left\{-cL(r-1) - \frac{1}{2c}L\left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right)\right\},\end{aligned}$$

so that the largest contribution is for $r = 2$. The function $c + 1/(2c)$ attains its minimal value $\sqrt{2}$ when $c = 1/\sqrt{2}$, hence

$$\mathcal{T}_{c,r}(x) \ll x\exp\left\{-\sqrt{2}L\left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right)\right\},$$

and we obtain

$$N(x) \le x\exp\left\{-\sqrt{2\log x\log_2 x}\left(1 + O\left(\frac{\log_3 x}{\log_2 x}\right)\right)\right\}. \tag{3.9}$$

This is somewhat weaker than the result implied by Theorem 1. But we saw that the main contribution to (3.7) comes from $\mathcal{T}_{c,2}(x)$, which is quite similar to the main sum appearing in the estimation (lower bound) of $T_0(x)$ (cf. [12]). The only difference is that in the case of $\mathcal{T}_{c,2}(x)$ we have $\Psi(xp^{-2}, 2p)$, while in the case of $T_0(x)$ the terms $\Psi(xp^{-2}, p)$ appear. In the relevant range for p (i.e., $\exp(\frac{1}{4}L) \le p \le \exp(2L)$) this will not make much difference, so that using the arguments [12] we can obtain for $N(x)$ the upper bound implied by the expression standing in (1.5)–(1.7). It is only for the sake of clearness of exposition that we gave all the details for the slightly weaker upper bound (3.9). This discussion completes the proof of Theorem 1.

4. Proof of Theorem 2

For the proof of Theorem 2 we need an expression that is sharper than (3.5), which is only an upper bound. If n is counted by $N(x)$ then

a) Either $P^2(n)$ divides n, and the contribution of such n is counted by $T_0(x)$ (cf. (3.1)), or

b) The number n is of the form $n = pq^bk, p = P(n), b \ge 2, P(k) < p, P(k) \ne q$, where henceforth q (as well as p) will denote primes. Since n does not divide $P(n)!$, then (similarly as in Section 3) it follows that $qb > P(n)$. Also, again as in Section 3, it is found that the contribution of $p \le \exp(L/4) = Y_1$ is negligible. The contribution of $p > \exp(2L) = Y_2$ is also negligible. Namely in this case, since $q^2|n, q > P(n)/b \gg \exp(2L)/\log x$, the contribution is clearly

$$\ll x \sum_{p \gg \exp(2L)/\log x} p^{-2} \ll x\exp(-\tfrac{3}{2}L).$$

Therefore, if b) holds, we need only consider numbers n for which there is a number $b \ge 2$ and a prime $q \in (p/b, p)$, such that $n = pq^bk, p = P(n), P(k) < p, P(k) \ne q$. The contribution of $P(k) = q$ can be easily seen to be negligible, as can be also shown for the one pertaining to $P(k) = p$. Using the technique of Section 3 (or of [8]) it is seen that the main contribution comes from the sum with $b = 2$. If $b = 2$, namely if $n = pq^2k$, $p/2 < q < p$ and $q_1^2|k$ with some prime $p/2 < q_1 < p$, then n is counted at least twice, but it is seen that the contribution of such n is $\ll x\exp(-\frac{3}{2}L)$. Therefore the contribution to $N(x)$ of integers satisfying b) equals

$$O(x\exp(-\tfrac{3}{2}L)) + \sum_{Y_1 < p \le Y_2} \sum_{p/2 < q < p} \Psi\left(\frac{x}{pq^2}, p\right). \tag{4.1}$$

This is different from Ford [8, eq. (2)], who had $\Psi\left(\frac{x}{pq^2}, q\right)$ instead of the correct $\Psi\left(\frac{x}{pq^2}, p\right)$. This oversight does affect his result (1.1), which is not correct since the constant is not the right one ($1 + \log 2$ should be replaced by 2), but the relation (1.3) remains valid, since both expressions with the Ψ-function are of the same order of magnitude. It follows that our starting relation for the proof of Theorem 2 takes the shape

$$N(x) = T_0(x) + O(x\exp(-\tfrac{3}{2}L)) + \sum_{Y_1 < p \le Y_2} \sum_{p/2 < q < p} \Psi\left(\frac{x}{pq^2}, p\right). \tag{4.2}$$

By using the asymptotic formula (2.3) for $\Psi(x, y)$, the prime number theorem in the standard form (see e.g., [14])

$$\pi(x) = \sum_{p \le x} 1 = \int_2^x \frac{\mathrm{d}t}{\log t} + O(x\exp(-\sqrt{\log x})), \tag{4.3}$$

we write the sum in (4.2) as a Stieltjes integral and integrate by parts. We set for brevity

$$R := \sqrt{\frac{\log_2 x}{\log x}}.$$

Then we see that the sum in question equals

$$\begin{aligned}&(1+O(R))x\int_{Y_1}^{Y_2}\frac{1}{\log t}\int_{t/2}^{t}\frac{1}{ty^2\log y}\rho\left(\frac{\log x}{\log t}-2\frac{\log y}{\log t}-1\right)\mathrm{d}y\,\mathrm{d}t\\ &=(1+O(R))x\int_{Y_1}^{Y_2}\frac{1}{t\log t}\int_{1-\log 2/\log t}^{1}\frac{1}{t^z z}\rho\left(\frac{\log x}{\log t}-2z-1\right)\mathrm{d}z\,\mathrm{d}t,\end{aligned}$$

on making the substitution $\log y = z\log t$. Since both z and $1/z$ equal $1+O(R)$ in the relevant range, by the use of (2.5) we see that our sum becomes

$$\begin{aligned}&(1+O(R))x\int_{Y_1}^{Y_2}\frac{\mathrm{e}^{2\xi(\log x/\log t-1)}}{t\log t}\rho\left(\frac{\log x}{\log t}-1\right)\int_{1-\log 2/\log t}^{1}t^{-z}\,\mathrm{d}z\,\mathrm{d}t\\ &=x(1+O(R))\int_{Y_1}^{Y_2}\frac{\mathrm{e}^{2\xi(\log x/\log t-1)}}{t^2\log^2 t}\rho\left(\frac{\log x}{\log t}-1\right)\mathrm{d}t.\end{aligned}\tag{4.4}$$

Now we use (2.6) (with $u-3$ replacing u, $v=-2$), setting $u=\log x/\log t$, $Y_1\le t\le Y_2$, to obtain that

$$\begin{aligned}\mathrm{e}^{2\xi(u-1)}\rho(u-1)&=\mathrm{e}^{2\xi(u-1)}\rho((u-3)+2)\\ &=\mathrm{e}^{2\xi(u-1)-2\xi(u-3)}\rho(u-3)\cdot(1+O(R)).\end{aligned}$$

Since $\xi'(u)\sim 1/u$ ($u\to\infty$) we have

$$\mathrm{e}^{2\xi(u-1)-2\xi(u-3)}=\mathrm{e}^{O(1/u)}=1+O\left(\frac{1}{u}\right)=1+O(R).$$

Therefore the last integral in (4.4) equals

$$(1+O(R))\int_{Y_1}^{Y_2}\rho\left(\frac{\log x}{\log t}-3\right)\frac{\mathrm{d}t}{t^2\log^2 t}.\tag{4.5}$$

With the change of variable

$$u=\frac{\log x}{\log t},\ \frac{\mathrm{d}t}{t\log^2 t}=-\frac{\mathrm{d}u}{\log x}$$

it follows that (4.5) becomes

$$
\begin{aligned}
S &:= \frac{1}{\log x}(1+O(R))\int_{y_1}^{y_2}\rho(u-3)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
y_1 &:= \tfrac12\sqrt{\frac{\log x}{\log_2 x}},\quad y_2 := 4\sqrt{\frac{\log x}{\log_2 x}}.
\end{aligned}
$$

But since $u\rho'(u) = -\rho(u-1)$, integrating by parts we obtain

$$
\begin{aligned}
S &= \frac{1}{\log x}(1+O(R))\int_{y_1}^{y_2} -u\rho'(u-2)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
&= \frac{1}{\log x}(1+O(R))\int_{y_1}^{y_2} \rho(u-2)\left(1+\frac{\log x}{u}\right)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
&= (1+O(R))\int_{y_1}^{y_2} -(u-1)\rho'(u-1)\left(\frac{1}{u}\right)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
&= (1+O(R))\int_{y_1}^{y_2} \rho(u-1)\left(\frac{\log x}{u^2}\right)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
&= \log x(1+O(R))\int_{y_1}^{y_2} -\rho'(u)\left(\frac{1}{u}\right)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,\\
&= \log x(1+O(R))\int_{y_1}^{y_2} \rho(u)\left(\frac{\log x}{u^3}\right)\exp\left(-\frac{\log x}{u}\right)\,\mathrm{d}u,
\end{aligned}
$$

where we used several times that $u \asymp 1/R$ in the range of integration, so that lower order terms could be absorbed by the $O(R)$–term. Making again the change of variable $\log x/\log t = u$, we obtain

$$
S = (1+O(R))\int_{Y_1}^{Y_2}\rho\left(\frac{\log x}{\log t}\right)\frac{\log t}{t^2}\,\mathrm{d}t = (1+O(R))\frac{T_0(x)}{x},
$$

where (3.4) was used. This proves (1.8).

5. Moments of the Smarandache function

In this section we shall prove Theorem 3. We begin with the proof of (1.9). Recall that the Smarandache function $S(n)$ denotes the smallest $k \in \mathbb{N}$ such that $n|k!$. This implies that $P(n) \le S(n) \le n$, and if $S(n) \ne P(n)$, then n does not divide $P(n)!$. Thus we may write

$$\begin{aligned}\sum_{2\le n\le x}\frac{1}{S^r(n)} &= \sum_{2\le n\le x, S(n)=P(n)}\frac{1}{S^r(n)} + \sum_{2\le n\le x, S(n)\ne P(n)}\frac{1}{S^r(n)} \\ &= \sum_{2\le n\le x}\frac{1}{P^r(n)} + O\left(\sum_{2\le n\le x, n\nmid P(n)!}\frac{1}{P^r(n)}.\right).\end{aligned} \tag{5.1}$$

We have, by Ivić–Pomerance [12],

$$\sum_{2\le n\le x}\frac{1}{P^r(n)} = x\exp\left\{-\sqrt{2r}L(x)\Big(1+g_{r-1}(x)+O\Big(\Big(\frac{\log_3 x}{\log_2 x}\Big)^2\Big)\Big)\right\}, \tag{5.2}$$

with $L(x)$ and $g_r(x)$ given by (1.6) and (1.7), respectively. On the other hand, by the argument that gives (3.5) and the proof of Theorem 1,

$$\begin{aligned}&\sum_{2\le n\le x, n\nmid P(n)!}\frac{1}{P^r(n)} \\ &\ll \sum_{2\le s\le \log x/\log 2}\ \sum_{p\le x^{1/s}}\frac{1}{p^r}\Psi\left(\frac{x}{p^s}, ps\right) \\ &\ll x\exp\left\{-(\sqrt{2(r+1)}+o(1))L(x)\right\} \qquad (x\to\infty).\end{aligned} \tag{5.3}$$

The proof of (1.9) follows then from (5.1)–(5.3), with the remark that a true asymptotic formula for the sum in (1.9) follows from [11, eq. (2.11)], since the major contribution to the sum in question comes from n for which $S(n) = P(n)$. However, similarly to the comparison between (1.5) and (1.8), the advantage of (1.9) is that the right-hand side contains only elementary functions.

The proof of (1.10) utilizes the following elementary lemmas, which generalize e.g., Lemma 1 and Lemma 2 of De Koninck–Ivić [4].

Lemma 1. *Let $r > -1$ be a fixed number, and let $c_{r,m} = (m-1)!(r+1)^{-m}$ for $m \in \mathbb{N}$. For any fixed integer $M \ge 1$ we have*

$$\sum_{p\le x} p^r = x^{r+1}\left(\frac{c_{r,1}}{\log x} + \ldots + \frac{c_{r,M}}{\log^M x} + O\left(\frac{1}{\log^{M+1} x}\right)\right). \tag{5.4}$$

Proof. Follows by partial summation from the prime number theorem in the form (4.3).

Lemma 2. *For a given fixed real number $r > 0$ and given natural numbers j and m, we have*

$$\sum_{n\le\frac12 x}\frac{1}{n^{r+1}(\log x/n)^j}=\sum_{k=0}^{m}\zeta^{(k)}(r+1)\binom{-j}{k}\frac{1}{(\log x)^{j+k}}+O\left(\frac{1}{\log^{j+m+1}x}\right). \tag{5.5}$$

Proof. Set $\ell=\ell(x)=\exp(\sqrt{\log x})$. Then

$$\sum_{n\le\frac12 x}\frac{1}{n^{r+1}(\log x/n)^j}=\sum_{n\le\ell}\frac{1}{n^{r+1}\log^j x\Big(1-\frac{\log n}{\log x}\Big)^j}+O(\ell^{-r}).$$

The assertion of the lemma follows if we use the binomial expansion

$$\left(1-\frac{\log n}{\log x}\right)^{-j}=\sum_{k=0}^{m}(-1)^k\binom{-j}{k}\left(\frac{\log n}{\log x}\right)^k+O\left(\left(\frac{\log n}{\log x}\right)^{m+1}\right)$$

and ($\zeta(s)$ denotes the Riemann zeta-function)

$$\sum_{n=1}^{\infty}\log^k n\cdot n^{-r-1}=(-1)^k\zeta^{(k)}(r+1)\qquad(r>0;k=0,1,\ldots).$$

Now we turn to the proof of (1.10). We have, by Theorem 1,

$$\begin{aligned}\sum_{2\le n\le x}S^r(n)&=\sum_{2\le n\le x,S(n)=P(n)}S^r(n)+\sum_{2\le n\le x,S(n)\ne P(n)}S^r(n)\\&=\sum_{2\le n\le x,S(n)=P(n)}P^r(n)+O\left(x^r\sum_{2\le n\le x,n\nmid P(n)!}1\right)\\&=\sum_{2\le n\le x}P^r(n)+O\Big\{x^{1+r}\exp(-(\sqrt2+o(1))L(x))\Big\}\quad(x\to\infty).\end{aligned}\tag{5.6}$$

Observing that $\Psi(x/p,p)=[x/p]$ if $\sqrt{x}\le p\le x$, it follows that $(k,n\in\mathbb{N},r>0)$

$$\begin{aligned}\sum_{2\le n\le x}P^r(n)&=\sum_{pk\le x,P(k)\le p}p^r=\sum_{p\le x}p^r\Psi\left(\frac{x}{p},p\right)\\&=\sum_{p\le\sqrt{x}}p^r\Psi\left(\frac{x}{p},p\right)+\sum_{\sqrt{x}<p\le x}p^r\Psi\left(\frac{x}{p},p\right)\\&=O\Big(x\sum_{p\le\sqrt{x}}p^{r-1}\Big)+\sum_{\sqrt{x}<p\le x}p^r\left[\frac{x}{p}\right]\\&=O(x^{1+r/2})+\sum_{pn\le x}p^r.\end{aligned}\tag{5.7}$$

By using Lemma 1 and Lemma 2 we infer that

$$
\begin{aligned}
\sum_{pn\le x} p^r &= \sum_{n\le \frac12 x}\ \sum_{p\le x/n} p^r \\
&= \sum_{n\le \frac12 x} \left(\frac{x}{n}\right)^{r+1} \left\{ \sum_{j=1}^{J} \frac{c_{r,j}}{\log^j (x/n)} + O\left(\frac{1}{\log^{J+1}(x/n)}\right)\right\} \\
&= x^{r+1}\sum_{j=1}^{J} c_{r,j} \sum_{n\le \frac12 x} \frac{1}{n^{r+1}\log^j (x/n)} + O\left(\frac{x^{r+1}}{\log^{J+1} x}\right) \\
&= x^{r+1}\sum_{j=1}^{J} c_{r,j} \sum_{i=0}^{J} \zeta^{(i)}(r+1)\binom{-i}{j}\frac{1}{(\log x)^{i+j}} + O\left(\frac{x^{r+1}}{\log^{J+1} x}\right) \\
&= x^{r+1}\sum_{k=1}^{J} \frac{a_{k,r}}{\log^k x} + O\left(\frac{x^{r+1}}{\log^{J+1} x}\right)
\end{aligned}
$$

with

$$
a_{k,r} := \sum_{i=0}^{k-1} (k-i-1)!(r+1)^{i-k}\zeta^{(i)}(r+1)\binom{-i}{k-i}. \tag{5.8}
$$

The assertion (1.10) of Theorem 3 follows then from (5.6)–(5.8), clearly one has $a_{1,r} = \zeta(r+1)/(r+1)$.

Finally we remark that, for $r \in \mathbb{R}$ fixed and $x \to \infty$,

$$
\sum_{2\le n\le x} \left(\frac{S(n)}{P(n)}\right)^r = x + O\left\{x \exp(-(\sqrt{2}+o(1)))L(x)\right\}. \tag{5.9}
$$

Namely, if $n = p_1^{\alpha_1}\dots p_k^{\alpha_k}$ with $p_1 = P(n)$ is the canonical decomposition of n, then n divides

$$
\{(\alpha_1+\alpha_2+\ldots+\alpha_k)p_1\}! \le \left\{P(n)\left(\left[\frac{\log n}{\log 2}\right]+1\right)\right\}!.
$$

Thus we have $P(n) \le S(n) \ll P(n)\log n$, and one obtains without difficulty (5.9) by the method of proof of Theorem 3.

References

[1] S. Akbik, *On a density problem of Erdős*, Int. J. Math. Sci. **22**(1999), No. 3, 655-658.

[2] R. de la Bretèche and G. Tenenbaum, Entiers friables: inégalité de Turán–Kubilius et applications, to appear, see `TKfriable.pdf` at `http://www.iecn.u-nancy.fr∼tenenb/`

[3] J.-M. De Koninck and N. Doyon, *On a thin set of integers involving the largest prime factor function*, Int. J. Math. Math. Sci. (2003), No. **19**, 1185-1192.

[4] J.-M. De Koninck and A. Ivić, *The distribution of the average prime factor of an integer*, Arch. Math. **43**(1984), 37-43.

[5] P. Erdős, *Problem* 6674. Amer. Math. Monthly **98** (1991), 965.

[6] P. Erdős, A. Ivić and C. Pomerance, *On sums of reciprocals involving the largest prime factor of an integer*, Glasnik Mat. **21**(41) (1986), 283-300.

[7] S.R. Finch, *Moments of the Smarandache function*, Smarandache Notions Journal **11**(2000), No. 1-3, 140-141.

[8] K. Ford, *The normal behavior of the Smarandache function*, Smarandache Notions Journal **10**(1999), 81-86.

[9] A. Hildebrand, *On the number of positive integers $\le x$ and free of prime factors $> y$*, J. Number Theory **22**(1986), No. 3, 289-307.

[10] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. Théorie Nombres Bordeaux **5**(1993), 411-484.

[11] A. Ivić, *On sums involving reciprocals of the largest prime factor of an integer II*, Acta Arith. **71**(1995), 229-251.

[12] A. Ivić and C. Pomerance, *Estimates for certain sums involving the largest prime factor of an integer*, in "Colloquia Math. Soc. J. Bolyai **34**. Topics in Analytic Number Theory", Budapest 1981, North-Holland, 769-789.

[13] I. Kastanas, *Solution to problem* 6674, Amer. Math. Monthly **98** (1994), 179.

[14] G. Tenenbaum, *Introduction to Analytic and Probabilstic Number Theory*, Cambridge Studies in Advanced Mathematics **46**, Cambridge University Press, Cambridge, 1995.

Aleksandar Ivić
Katedra Matematike RGF-a
Universitet u Beogradu
Đušina 7, 11000 Beograd, Serbia
`aivic@rgf.bg.ac.yu`